Description of the MHS Health Level 7 Pharmacy Outpatient Data for Public Health Surveillance

Technical Document NMCPHC-EDC-TD-2-2014

By Ashleigh McCabe
EpiData Center Department
October 2009

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
30-10-2009	Technical Document	October 2006-October 2009

4. TITLE AND SUBTITLE
Description of the MHS Health Level 7 Pharmacy Outpatient Data for Public Health Surveillance

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Ashleigh McCabe

5d. PROJECT NUMBER
XXXXXX

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Navy and Marine Corps Public Health Center
EpiData Center Department
620 John Paul Jones Circle, Suite 1100
Portsmouth, VA 23708

8. PERFORMING ORGANIZATION REPORT NUMBER
NMCPHC-EDC-TD-2-2014

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Navy and Marine Corps Public Health Center
EpiData Center Department
620 John Paul Jones Circle, Suite 1100
Portsmouth, VA 23708

10. SPONSOR/MONITOR'S ACRONYM(S)
NMCPHC-EDC

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
NMCPHC-EDC-TD-2-2014

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release. Distribution is not limited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The EpiData Center Department (EDC) at the Navy and Marine Corps Public Health Center evaluated the Health Level 7 (HL7) data source for its usefulness in health surveillance activities. This technical document provides a history of the HL7 pharmacy outpatient database and its contents, explains the creation of prescription records, describes the pathway of data from healthcare provider to the EDC, provides a detailed descriptions of all variables within the database, and assesses the database's strengths and limitations. Given an understanding of the strengths and limitations of the data, HL7 pharmacy outpatient data have proven to be a valuable source of health information for surveillance purposes. The data can be used for case identification when disease-specific treatment is available, or used to assess clinical practice guideline adherence for known cases. Furthermore, data are received in a timely fashion, allowing for near-real-time surveillance of diseases.

15. SUBJECT TERMS
Health Level 7 (HL7), Pharmacy, Prescription, Surveillance

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	25	757-953-0692

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Abstract

The EpiData Center Department (EDC) at the Navy and Marine Corps Public Health Center evaluated the Health Level 7 (HL7) data source for its usefulness in health surveillance activities. This technical document provides a history of the HL7 pharmacy outpatient database and its contents, explains the creation of prescription records, describes the pathway of data from healthcare provider to the EDC, provides a detailed descriptions of all variables within the database, and assesses the database's strengths and limitations. Given an understanding of the strengths and limitations of the data, HL7 pharmacy outpatient data have proven to be a valuable source of health information for surveillance purposes. The data can be used for case identification when disease-specific treatment is available, or used to assess clinical practice guideline adherence for known cases. Furthermore, data are received in a timely fashion, allowing for near-real-time surveillance of diseases.

Table of Contents

Executive Summary

Project Background

The EpiData Center (EDC) at the Navy and Marine Corps Public Health Center (NMCPHC) was funded to evaluate the Health Level 7 (HL7) Pharmacy Outpatient (OP) data source for its usefulness in public health surveillance. This technical document is the result of the analysis of the Pharmacy OP data which is one of three Pharmacy data types: Outpatient, Unit Dose, and Intravenous. The OP dataset contains the pharmacy records associated with ambulatory and outpatient visits dating since 01 Oct 2006.

Public Health Surveillance Applications

OP data add a unique layer to the EDC's surveillance efforts. Because these data are not limited to laboratory confirmed cases, they can provide information on presumptively treated cases, which are common because of over-burdened facilities and limited treatment options. Also, where the treatment of a disease uses a specific medication, these data can indicate the diagnosis more precisely than the diagnosis codes from inpatient or encounter records, since the codes in these records may be imprecise. Data on pharmacy transactions, therefore, can improve the robustness of surveillance systems based on lab results and/or clinical encounters.

Key Fields for Public Health Surveillance

Specific key fields for Public Health Surveillance are included in the data: National Drug Code (NDC) code, NDC number, Sponsor ID, Family Member Prefix (FMP), Service, Requesting Facility, and Performing Facility.

True duplicates are defined as records in which all fields are exactly the same. After true duplicates are eliminated, the data can be analyzed by unique prescription order, individual, or medication. Unique patients may be identified in the HL7 pharmacy data through a combination of sponsor ID and FMP, this combination creates a unique identifier that can be used to track individual patients through all pharmacy records. A unique order may be defined as all records associated with each specific drug prescription. A unique record may be defined as all transactions associated with each prescription for an individual patient.

There are several fields that relate directly to the medication dispensed: Amount Dispensed, Drug Name, NDC Code, Provider’s Administration Instructions and Units. These fields have several unique characteristics that should be considered prior to analysis.

Strengths

Several of the data fields of interest are complete but the completeness of the database as a whole continues to be assessed. The analysis showed that data were collected in the Composite Health Care System (CHCS) from the majority of the Department of Defense (DOD) Medical Treatment Facilities (MTFs). The timeliness of reporting is within the acceptable range for the Navy surveillance activities, 1-3 days.

Limitations

It is currently not clear whether Defense Health Services Systems (DHSS) captures all CHCS pharmacy transactions. Further work is needed to compare HL7 pharmacy records to other data sources in order to estimate completeness. The OP data only includes MTFs that have CHCS servers. Therefore, forward deployed clinics, contracted managed care support clinics, and other MTFs that do not use CHCS are not captured in these data unless the prescription is taken to an MTF to be filled at a pharmacy that uses CHCS. Incomplete demographic information (e.g., marital status, race, and ethnicity) can limit the generalizability of these data to specific minority groups. Extra precautions need to be taken when extrapolating data to larger populations and when comparing disease rates and trends among the military to non-military populations.

Introduction

The EpiData Center (EDC) at the Navy and Marine Corps Public Health Center (NMCPHC) was funded to evaluate the Health Level 7 (HL7) Pharmacy data source for its usefulness in public health surveillance. This technical document resulted from the analysis of the Pharmacy Outpatient data string which is one of three Pharmacy data types: Outpatient, Unit Dose, and Intravenous. The Outpatient Pharmacy (OP) dataset records non-intravenous prescriptions filled at a military treatment facility (MTF) for prescriptions generated during an ambulatory (outpatient) visit. Records for all Department of Defense (DOD) military service members (Army, Navy, Marine Corps, Air Force, Coast Guard, US Public Health Service), overseas civilian personnel, Tri-Care eligible dependents, and others who receive their prescriptions at a military MTF are included in this data string. The following document describes the original observations on the data fields, some statistics, the comments received from Defense Health Services Systems (DHSS) program office of Tri-Care, DOD partners or other staff, the modifications made, the cleaning rules implemented for usability, and other comments relevant to the use of these data for surveillance.

Initial evaluation of the data stream involved three sample OP extracts received by the EDC from DHSS. These were small datasets used to analyze the structure, completeness and distribution of the entire data set. The first OP extract was received for the Message Date of 07/29/2006 (extract I), and a second extract was generated for 09/14/2006-09/16/2006 (extract II). Descriptive analysis on these data included frequency distribution of demographic fields, evaluation of null or invalid values for key fields used in surveillance, and understanding data structure in the extracts received compared to the structure as data is entered into the Composite Health Care System (CHCS). A third extract, which covered the dates 02/07/2006-02/09/2006 (extract III) was used to assess more accurately the completeness and the usefulness of these methods in ongoing influenza work at the EDC. All three extracts were reviewed and analyzed in order to determine applicability for surveillance and modify the data structure to more accurately address the disease surveillance needs of the Navy and Marine Corps. After the initial analysis demonstrated the potential value of the data for surveillance activities, the EDC began receiving daily feeds. The current data archive dates back to 01 Oct 2006.

Data Origination and Flow Process

The pharmacy OP data stream includes all prescriptions that are filled at an MTF pharmacy. There are several mechanisms of entry that can occur. The most common process followed is described below along with notable exceptions.

A medication order is initially entered into the CHCS system by the prescribing (requesting) provider. The patient then reports to the pharmacy to pick up the prescription. The pharmacist receives the order via CHCS and verifies it. When the pharmacist fills the order and dispenses the medication, he or she completes the record and saves it in the local CHCS system. If a prescription is edited upon verification, edits are made in the CHCS record. The pharmacist has the ability to cancel prescriptions per the physician or when the medication is not picked up by the patient. Each time a record is canceled, changed, edited, reordered, or refilled a new record in CHCS is generated.

An alternative record creation process in CHCS is used when the prescription is received from a non-CHCS participating provider and filled at an MTF pharmacy. In this case, a written prescription is submitted to the pharmacy. The pharmacist creates the order and completes the record in CHCS as the prescription is verified and filled. This process can occur in several circumstances, including non-MTF doctor's visits and prescriptions written by ship-based clinicians, as neither use the CHCS system for ordering medications. All these activities in CHCS also generate HL7 messages for that prescription. Specifically, an HL7 message is generated when a label is printed for a new, refilled, or edited prescription. An HL7 message is also generated when a prescription is entered manually at a site that does not print labels for manual prescriptions, is marked non-compliant in CHCS, or is removed. Edited, refilled, and cancelled prescriptions in the HL7 data will have the same order number as the original HL7 message for that prescription.

Dental clinics are associated with an MTF or ship, and prescriptions written there, like those written at an outpatient clinic, are received from the parent facility's pharmacy. Therefore the records follow the same entry pattern as those of their parent facility. Depending on the initial CHCS set up, the clinic may not be explicitly named in the requesting facility fields. In this case, the dental clinic's parent facility may be listed in the requesting facility name. Consequently, two possible fields for identifying records from a dental clinic are Requesting Work Center and Medical Expense and Performance Reporting System (MEPRS) code.

The HL7 Pharmacy data are limited to prescriptions filled at an MTF pharmacy that uses CHCS. If prescription orders are entered into CHCS and not filled (a label is not printed at the pharmacy), these medications are not seen in the HL7 pharmacy OP data stream. Prescriptions filled in a network pharmacy are also not included in this data stream. However, HL7 is not the only source for pharmacy data in the Military Health System (MHS). The Pharmacy Data Transaction Service (PDTS) is a centralized data repository that collects prescription information for all DOD beneficiaries that are filled at MTFs, retail locations and mail order pharmacies. This

service is set up as real time provider of transaction support to ensure patient safety. During a patient visit, when a provider enters a prescription order into CHCS, information is sent to the PDTS to review medication history and ensure that the medications being prescribe will not adversely interact with other medications the patient is currently taking. The provider receives a response on his/her CHCS screen within 15 seconds. PDTS data are collected, therefore, almost instantaneously and include information on prescription orders as opposed to medication fills. The Pharmacoeconomic Center has done extensive work with the PDTS data and continues to support Tri-Care decision making (including formulary set ups) using these data. Though these data include more prescriptions to beneficiaries that were not filled at military MTFs, they do not include inpatient medication transactions, whereas the three HL7 pharmacy data streams include both outpatient and inpatient transactions but lack data on prescriptions not filled at MTFs. PDTS data are not currently available to the EDC.

Public Health Surveillance Applications

OP data add a unique layer to the EDC's surveillance efforts. Because these data are not limited to laboratory confirmed cases, they can provide information on presumptively treated cases, which are common because of over-burdened facilities and limited treatment options. Also, where the treatment of a disease uses a specific medication, these data can indicate the diagnosis more precisely than the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes from inpatient or encounter records, since the codes in these records may be imprecise. Data on pharmacy transactions, therefore, can improve the robustness of surveillance systems based on lab results and/or ICD-9-CM coded records.

The greatest value of pharmacy data for the Navy and Marine Corps currently lies in disease-specific treatments. However, many symptoms and treatments are not specific to a particular disease or condition. It is therefore necessary to fully understand the treatments for a disease of interest and be aware of the other indications for which those treatments may be used. Treatments for conditions such as influenza, malaria, and tuberculosis are relatively specific and may be useful proxies for a diagnosis when the dosage and length of treatment are considered.

Current surveillance methods in the EDC include monitoring HL7 lab results and ICD-9-CM codes or encounters recorded in the Standard Ambulatory and Inpatient Data Record (SADR/SIDR). Consequently, surveillance methods are largely disease-specific, but this specificity depends on ICD-9-CM coding practices in local MTFs, the timeliness of lab testing, and the ability to accurately flag lab tests of interest. The use of pharmacy data will greatly improve the surveillance of certain infectious diseases, such as tuberculosis, because other data on these diseases are so limited by delayed lab test results and gross inaccuracies in ICD-9-CM coding.

Potential use of HL7 pharmacy records is not limited to surveillance. Data on dispensed medications can fill critical gaps in the military's ability to track medication compliance with regard to outcomes such as treatment of latent tuberculosis infection, high blood pressure, diabetes, or sexually transmitted diseases. Coupled with laboratory and encounter data, disease management guidelines can be evaluated. Finally, these data may provide valuable insight into antibiotic therapy and subsequent emerging resistance. An example of analysis completed using the HL7 pharmacy outpatient dataset is provided in Appendix A.

Data Structure and Analysis

HL7 Pharmacy data are retrieved by the EDC in a standard, pipe-delimited flat file from DHSS. Each column within the data file is a fixed variable and each row should contain a unique record. Each person can have more than one distinct record, if they have multiple medications, edits, reorders, or cancellations. Each medication prescribed is associated with a unique record (row). Any additional changes (edits, cancellations, refills, etc.) to that record are in a separate record than the original, but have the same order number. Variable fields are designed to ease analysis, except for the free text fields, which will require the use of wildcards or search terms.

Key Fields for Public Health Surveillance

Defining Duplicates

Within the HL7 Pharmacy data string there are several ways in which unique records can be identified. To date the EDC has only used Pharmacy data in conjunction with influenza surveillance. Duplicate rules should be checked against other project objectives to ensure the desired outcome results. True duplicates are defined as records in which all fields are exactly the same. Records meeting this criterion should be eliminated. After true duplicates are eliminated, analysis should take into account that any changes to a new or existing medication transaction appear as a separate record in the HL7 database. As described below, the data can be analyzed by unique prescription order, individual, or medication

Unique ID/Case

Patients are identified in the HL7 pharmacy data through a combination of Sponsor ID and FMP. This combination creates a unique identifier that can be used to track individual patients through all pharmacy records. It is important to note that it is possible for individuals to have two separate Sponsor IDs over time. For example, if the spouse of a sponsor becomes active duty, then that spouse will have his/her own Sponsor ID. Each unique patient can have multiple medication orders in the HL7 pharmacy data. A unique order is defined as all records associated with a single specific drug prescription (one medication). Each medication prescribed and filled is assigned its own unique order number. The combination of Sponsor ID, FMP, and Order Number creates an additional unique identifier at the level of order. Using this unique order identifier, the analyst will be able to identify an original order and all changes and modifications made to that order, such as dosage and administration order changes. It is important to note that the data do not contain all medications that were ordered, only those medications for which a label was printed in the pharmacy. Each unique order can have multiple records within the HL7 pharmacy data. A unique record is defined as an individual transaction associated with each prescription for an individual patient (new order, edited order, or cancelled order). The combination of Sponsor ID, FMP, Order Number, and Order Control creates a final unique identifier that matches each completed transaction to a specific prescription order. One unique order may have multiple unique records depending on any changes in the transaction. The values of Order Control represent the status of the prescription, and it is often necessary for cancelled, edited, replaced, and unknown status records to be removed prior to analysis.

Medication Dispensed

There are several fields that relate directly to the medication dispensed: Amount Dispensed, Drug Name, National Drug Code (NDC), Provider's Administration Instructions and Units. These fields have several unique characteristics that should be considered prior to analysis.

The Drug Name field should be standardized by the analyst. While this field at the minimum contains the drug name, it varies widely in terms of content. It is often necessary to search for the drug name of interest and to remove other extraneous information contained in this field.

For example, the drug acetaminophen is present in the following forms for this field: ACETAMINOPHEN (CHILDREN'S TYLENOL) 160M, ACETAMINOPHEN 325MG ORAL TABLET and ACETAMINOPHEN 500MG ORAL TABLET.

Provider's Administration Instructions are the actual written directions by the clinician, and are not formatted for ease of analysis. This free text field could be searched for terms relating to a specific prescription if deemed necessary.

Dosage of a medication is an important aspect of study design for analysis of these data. Due to clinical practice, it is common to see the same medication used for prophylaxis and for treatment in the HL7 data. In order to distinguish the two situations, the analyst must know the appropriate dosage of the medication for each situation and the dosing schedule. It is important to consider whether the medication and dosage of interest is disease or condition specific. For example, amantadine is an anti-viral used to treat influenza but is also used to control body spasms associated with Parkinson's disease.

Strengths

Timeliness

DHSS includes several date and time fields in the data string provided to the EDC: Message Date, DHSS Load Date, Transaction Date, and Date of Most Recent Refill. A timeline of dates is provided in Appendix E. To assess the timeliness of the data, the Transaction Date (date the order was placed into CHCS by the provider) was compared to the Message Date (date the HL7 message was generated by CHCS) to estimate the time between the patient encounter and the receipt of data at DHSS. The Message Date was also compared to the DHSS Load Date in order to determine the time between HL7 message generation at the local CHCS host and DHSS data parsing of the HL7 message into the database design.

On average, it took less than a day for an HL7 message to be generated. After generation, it took about 1 day for the message to be processed by DHSS with a range of 0-11 days (the most frequent lag times are 0, 1, or 2 days). It is assumed that NMCPHC receives these data within two days, though this assumption should be verified in the future as time permits. This interval indicates that the timeliness of reporting is within acceptable ranges for the Navy surveillance activities. Future analysis and assessment should better define lag times in relation to particular MTFs, drugs or disease outcomes of interest.

Completeness

Records are received from the majority of shore-based MTFs of the CHCS system. As described in the limitations section below, data transmission depends on a pilot network set up by DHSS. Due to the limitations of the network, gaps in the data do exist. Data was not received for two small periods of time, and therefore might be missing or incomplete for message dates 12/06/2006-12/07/2006 and 03/11/2007-03/12/2007. Even so, it is believed that these data may represent at least 90% of all filled prescriptions in CHCS. Further analysis is being conducted to quantify completeness.

The completeness of individual fields varies and the characteristics of each are described in detail in the field observations section that appears later in this document. In general, some fields of particular interest, such as Sponsor ID, FMP, and Service are highly populated due to the business rules of CHCS.

Limitations

Completeness

Previous DHSS studies have evaluated DHSS's capture of pharmacy transactions from the local MTF. However, these studies were limited to one MTF and one day. It is currently not clear whether DHSS captures all CHCS pharmacy transactions. Further work is needed to compare HL7 pharmacy records to other data sources in order to estimate completeness.

The HL7 infrastructure at DHSS was built using pilot funds. The network was created as a temporary network that captures HL7 messages when they are sent from the CHCS host to the Denver feed node. Currently, no back-up system exists. Consequently, when the feed node goes down, HL7 messages are lost and those that have been sent cannot be retrieved unless the network outage was planned for in advance. Therefore, gaps exist in the data received at NMCPHC. Analysis is being conducted to estimate the data loss and its impact on surveillance activities. Several of the identified data fields of Public Health interest are highly populated, but other fields included in the data are not reported as completely. The completeness of each data field, as described later in this document, should be considered before use in analysis.

Inclusion

The OP data only includes MTFs that have CHCS servers. Forward deployed clinics, contracted managed care support clinics, and other MTFs that do not use CHCS are not captured in these data unless the prescription is taken to an MTF to be filled at a pharmacy that uses CHCS. The CHCS system is not used to order or fill prescription medication on board ship. If shipboard personnel are referred to shore clinics or pharmacies for medications, this information is captured in the HL7 data. In addition, dental clinics do not routinely appear in these data, though there are a small number of records that are reported from dental clinics. Further inquiries are required to understand why they do not appear. It is possible that these are outlying clinics with no CHCS connection.

Generalizability

Incomplete demographic information (e.g. marital status, race, and ethnicity) can limit the generalizability of these data to specific minority groups. Demographic information not provided in this database can be supplemented with other available personnel databases.

Comparability

These data are generated from the pharmaceutical treatment records of a highly specific patient population – military service members and other military beneficiaries – which differs from the general U.S. population in many ways, including average age, gender distribution, physical fitness, and health status. Further, this population has universal access to medical care, which is not true of many people living in the U.S. These differences limit the comparability to the general US population. Extra precautions need to be taken when extrapolating data to larger populations and also when comparing the disease rates and trends of the military and non-military populations.

All Data Fields (Variables)

The following section describes frequency distributions run on all fields for each extract associated with the HL7 Outpatient data. Any problems that arose in relation to data values were addressed with DHSS and resolved to the best possible conclusion. The data fields of most interest include. NDC Code, NDC Name, Sponsor ID, FMP, Service, Requesting Facility, Performing Facility, and other such fields that are necessary for the EDC's planned surveillance activities.

Automatically Populated Fields

There are several types of automatically populated fields in the Pharmacy OP data.
When a facility registers within the CHCS system, several variables are created, which identify the facility: Performing DMIS ID, Performing Facility, Performing Facility Service, Performing Work Center, Pharmacy Site, Requesting DMIS ID, Requesting Facility, Requesting Facility Service and Requesting Work Center. When DHSS compiles the data from the CHCS server, two fields are automatically populated: DHSS Load Date and DHSS Load Time.

Each patient or beneficiary is registered in the Defense Eligibility Enrollment Reporting System (DEERS) under the Sponsor ID, which feeds into the CHCS system. When a patient presents at a medical facility the Sponsor ID (usually the Social Security Number) is entered and their name is chosen from a drop down list. The following patient demographic fields are automatically populated after this selection, if they were entered when the patient was registered in DEERS: Date of Birth, Ethnicity, FMP, Gender, Marital Status, Patient Category, Patient ID, Race, Service, Sponsor ID, Sponsor UIC Code and Sponsor UIC Description. If these data are not present in the system, a designated unknown value is entered, and therefore there are no missing values in these fields. Registration is completed and records updated when the sponsor reports to a new UIC and selects an MTF. Administrative personnel at the MTF have the ability to edit records at the time of visit.

As records are created, edited, and completed several variables are created by the CHCS system: Date of Most Recent Refill, Date of Transaction, Time of Most Recent Refill and Time of Transaction. These can be changed, if necessary, by the pharmacist, but this change is not common practice. MSG Date, MSG Time, and MSG Sending Facility are created and assigned when the message (record) is sent to the CHCS server. A timeline of the most common order of date variables observed in the data is presented in Appendix B.

Field Observations (in alphabetical order):

Appendix E contains a timeline and a brief description of the date variables in the data set. This appendix should be consulted to better understand the relationship of the dates included.

Amount Dispensed

The Amount Dispensed field indicates how much of the medication is dispensed. It can contain the number of pills, number of milliliters, number of tubes, etc. The units of this value are indicated in the Units field.

The Amount Dispensed field is important in identifying the purpose of the prescription. For instance, recruit trainees are not allowed to possess medications during their training programs. Because of this regulation, the medical corpsman assigned to them is responsible for obtaining all of their medications, including any prophylaxes. For example, if a person were identified by the pharmacy data as receiving 200 doses of penicillin, an analyst would need to know that that prescription may not be for the personal use of that patient. The possibility that medications are linked to a medical corpsman rather than the patient should be considered during analysis.

Date of Birth

The Date of Birth field is formatted YYYYMMDD. Birthdates were valid in the extracts examined.

Date of Most Recent Refill

This field indicates the date when the prescription was most recently refilled. It is only populated for those medications indicating refills are available and is formatted YYYYMMDD. The final message contains the refill date and time, if applicable. Records with an Order Control of RP (replacement order) do not have a value in this field. Replacement orders are messages sent through CHCS that indicate a previously ordered medication should not be processed because it will be replaced by a new prescription. Another message with an Order Control of NW (new) should follow the RP order.

This field can be useful in monitoring treatment compliance with long term medications such as the treatment for tuberculosis or maintenance medications used for chronic conditions. For example, the dosage of medication dispensed and the average time it takes for the patient to refill the prescription may determine if the patient is adhering to the physician's treatment instructions. This field can also help to determine if the medication had been dispensed to this patient previously without having to search the database.

Date of Transaction

The Transaction Date is the date on which the order entered the CHCS system. It is different from the Message Date since the Message Date is generated when the label is printed and the Transaction Date more accurately approximates when the prescription is actually presented to the pharmacy. The Transaction Date is formatted YYYYMMDD and does not include missing values.

Since this field approximates the time that the medication is ordered it may be useful for future analysis. It can be used to identify medications that are pre-ordered (for surgeries), and can allow analysts to compare the dates when a prescription is issued and picked up, and to assess the lag between the ordering of a prescription and the availability of its data to the EDC.

Drug Name

The Drug Name field is a text translation of the NDC code. The general format is: Scientific Name (Trade Name) Dosage. However, this format is not consistent. In many cases the last portion, after the trade name, is cut off at various places. This truncation creates problems when trying to analyze the frequency of medications using the drug name field. An alternative to this method would be to use the NDC Code field and then match to the corresponding drug name from another source. Those records missing a value for the drug name field are also missing the NDC code.

Records with some drug names are filtered out of the pharmacy data on a regular basis. The filter excludes the full records for all medications specifically used to treat HIV/AIDS. HIV/AIDS medications are identified from the Tri-Care formulary under HIV specific anti-retrovirals.

DHSS Load Date

DHSS Load Date indicates the time when DHSS prepares and sends the data to the EDC. The field is used to determine the timeliness of reporting and to identify lags in reporting times from certain MTFs. The format is YYYYMMDD.

DHSS Load Time

Time component of the DHSS Load Date field, and is formatted: HHMM.

Ethnicity

Ethnicity is an alpha numeric field with six possible values: 1=Hispanic, 2=South East Asian, 3= Filipino, 4=Other Asian Pacific Islander, 9=Other, and Z=Unknown.

There are no missing values in this field. However, approximately 96% of records in extracts examined indicated either unknown or other. These results indicate that the Ethnicity field may be self-identified and is not consistently reported. This lack of specificity potentially limits the ability to describe disease trends and burdens in minority groups unless the data are supplemented from other personnel information.

FMP

FMP is the Family Member Prefix, which designates the relationship of the patient to the sponsor. The distribution of FMP among the records for extracts II and III is as expected. The most common entries in the FMP field were 1-3, 20, and 30. These codes indicate a first, second, or third child of a sponsor, the sponsor, and the sponsor's spouse, respectively. A full list of FMP codes is available for analysis.

FMP is automatically populated in the CHCS system. This process occurs when the Sponsor ID is entered and the patient name is selected from a list of options, each of which is attached to a specific FMP.

Gender

There are three values possible for the Gender field; M=Male, F=Female, X=Unknown.

There are no missing values for gender and there is less than 1% with a value of unknown.

Marital Status

There are nine values for Marital Status: A=Annulled, D=Divorced, I=Interlocutory Decree, L=Legally Separated, M=Married, N=Never Married, S=Single/Not Married, W=Widow or Widower, Z=Unknown.

There are no missing values for Marital Status in the extracts examined. However, approximately half of all records are categorized as unknown. The Single/Not Married and Married categories account for the largest percentages of those with known values.

MEPRS Code

The Medical Expense and Performance Reporting System (MEPRS) code is a four letter code that indicates where within the MTF the person received treatment. The first letter indicates the most general area and translates as: A=inpatient, B=outpatient, C=Dental, D=ancillary, E=support, F=special programs, G=readiness. It is advised to obtain an up-to-date list of all possible codes. The field is useful for tracking where people are seen within the MTF. For instance, it can indicate ambulatory care, special dialysis clinics, the maternity ward, etc., which can affect the interpretation of the data.

The majority of records contained in the OP dataset will have an MPERS that begins with B. There are several inpatient records (MEPRS=A) included in the Outpatient data set, which may pertain to cases where a patient was admitted to the hospital but picked up medications after discharge.

Message Date

This field is automatically assigned in CHCS when the order is completed in the system and sent to the CHCS server. The date approximates the transaction date but this approximation can vary by location. Some MTFs send messages in batches, therefore the date portions may not correlate to the actual transaction date. This field is formatted: YYYYMMDD. There are no missing values for this variable.

Message Time

This field is automatically assigned in CHCS when the order is completed in the system and sent to the CHCS server. The time approximates the transaction time but this approximation can vary by location. Some MTFs send messages in batches, therefore the time portions may not correlate to the actual transaction time. This field is formatted: HHMM. There are no missing values for this variable.

Message ID

Message ID is an alphanumeric code assigned to each batch of messages based on when the message is sent from CHCS to the server. The Message ID is not unique to each record; each batch of messages is assigned one Message ID. The format of Message ID varies by MTF and

includes numbers, letters and/or a numeric code that identifies the MTF or it can identify the function of the message (e.g. RESCHED-057342).

Message Sending Facility

This field identifies the facility that sends the message to DHSS through the CHCS system. This field allows analysts to identify and track problems that arise in the transfer of messages from the MTFs through DHSS to the EDC.

NDC Code

The National Drug Code (NDC) is a unique three-segment code used to identify a drug. The segments are separated by "\". Every drug manufactured, prepared, propagated, compounded, or processed for commercial distribution is required to be registered with the FDA (Food and Drug Administration) and receives an NDC. The first segment of the NDC is the labeler code, which identifies the company that manufactures or distributes the drug under its label. The second portion of the NDC is the product code which identifies the strength, dosage and formulation of the medication. The final segment is the package code which identifies the package size and type (i.e. number of pills). The NDC can have several configurations of character lengths: 4-4-2, 5-3-2, or 5-4-1.

Number of Refills Remaining

This field is the Number of Refills remaining on the original prescription. There are only nine records from extracts II and III that have a missing value for this field. By far, the most frequent value is 0 which can indicate that all of the refills are used or that no refills are indicated on the original prescription. This field could be used to track a patient through their entire treatment regimen by following the decreasing value in this field. It will also allow an analyst to see if there are missing records. For instance, if there were a record indicating three refills remaining followed by a record indicating one refill remaining, further investigation could locate the missing record.

Order Control

The Order Control field designates the type of order: CA=cancel, HD=hold, NW=new, RE=refill, RL=release, RN=renew, RO=replacement order, RP=replace order-modify, XX=edited order, and ZP=prevalidate-preverification request and response. This field allows analysts to track some of the changes made to an order over time, as well as to distinguish refills from new prescriptions. The identification of the order type is important if the question of interest relates to incident cases.

A value of ZC is present in the extracts but is not identified as a possible value. The data dictionary also identified NC=non-compliant as a possible value but it is not included in any of the extracts.

Order Number

Order number is a numeric code of eleven digits (xxxxxx-xxxxx) unique to each order but not unique for each record. An order can have multiple records that correspond to changes made

to the order (e.g. changes in dosage or frequency of application, cancellations). All changes appear as individual records with the same order number. It is a plausible way to track a patient but it is not useful for identifying unique records.
There are no missing values in the order number field.

Ordering Provider

The Ordering Provider field indicates the name of the prescribing physician. It has four components separated by "\": Last Name\First Name\Middle Initial\Suffix. It is structured to facilitate analysis but could be separated if necessary.

Patient Category

The Patient Category Code (PatCat) is an alphanumeric code that indicates the patient's relationship to the uniformed services. The first letter of the code refers to the branch of service of the Sponsor (A=Army, B=National Oceanic and Atmospheric Administration, C=Coast Guard, F=Air Force, K=other beneficiaries of the federal government, M=Marine Corps, N=Navy, P=US Public Health Service, R=NATO Recipient). It is followed by two digits corresponding to the status of the Sponsor, as well as the patient's relationship. For example: A11=Army Active Duty Member, A41=Army Dependant of Active Duty, etc. A complete list is available from the DOD for analysis. The most frequently reported sponsor portion of the PatCat codes in all branches are 11 (active duty), 12 (reserve) and 13 (recruit).

Patient ID

The Patient ID is intended to serve as a unique identifier for each patient. In the data dictionary provided by DHSS, it states that the Patient ID is the Patient's SSN, when available. The use of a character type field preserves leading zeros that might otherwise be dropped.

Based on the EDC evaluation, the Patient ID should not be used to identify patients. Patient ID is often missing and may be a facility ID number assigned to a patient instead of their SSN. It is recommended that Sponsor ID and FMP be used together to create a unique identifier instead.

Performing DMIS Facility Name

This field is the text translation of the DMIS ID provided in the Performing DMIS ID field. This field is assigned by DHSS at the request of the EDC. The translation of the DMIS code on the official list is often more accurate than the Performing Facility field in CHCS. Therefore, it allows for more accurate investigations when geographic information is used. Since the field is also a translation of the Performing Facility field in CHCS, it is missing when that field is missing in the record.

Performing DMIS ID

The Performing DMIS ID is a four digit. The DMIS code is assigned by the Department of Defense to all units within all of the installations in order to uniquely identify them. The EDC provides an official DMIS list to DHSS for this purpose. DHSS translates the Requesting Facility field within CHCS to an assigned DMIS code. This code allows for grouping of MTFs based on geographic location, as well the ability to identify parent/child relationships between

installations. Since this field is calculated based on the Performing Facility field, all records missing a value for that field are missing a value for the Performing DMIS ID field. Since several of the codes begin with zeros they are often dropped when importing the data from a text file. This can lead to problems when trying to produce summary statistics based on MTF. Use of Character type field preserves leading zeros that might otherwise be dropped.

Performing Facility

The Performing Facility field in CHCS indicates the name of the MTF where the order is filled. Problems may be encountered if the text is entered incorrectly when the facility registers in the system (e.g. misspellings). This field allows for the tracking of orders from their origin to where they are filled. The translation of the DMIS code on the official DOD list in Performing DMIS Facility Name is often more accurate than the Performing Facility field in CHCS.

Performing Facility Service

The Performing Facility Service field indicates the branch of service with which an MTF is associated. This value is determined from the DMIS code list provided to DHSS by the EDC. It is missing from a record when the Performing Facility information is missing. The possible values are: A=Army, C=Coast Guard, F=Air Force, N=Navy. This field is useful for stratifying the observations included in any investigation. Often, the data available for use are only available for one service branch, that of the MTF or the patient. If this is the case, the HL7 pharmacy data are limited to the same parameters.

Performing Work Center

The Performing Work Center field indicates the specific work location within the pharmacy or MTF that provides the service. The field is still present in daily feeds but is unpopulated.

Pharmacy Site

The Pharmacy Site field indicates which pharmacy prepared the prescription. This is a text field that describes the location type of the pharmacy (e.g. SATELLITE PHARMACY or ER AFTER HOURS PHARMACY).

Provider's Administration Instructions

The Provider's Administration Instructions field is an unstructured text field. It provides the physician's instructions for how to use the medication (e.g., Take two times a day, Apply as directed by physician, etc). This field may be too cumbersome to use in direct analysis but it could be used to support findings from other fields. It can provide valuable information on dosing schedules and help indicate the purpose for which the medication is prescribed.

Race

There are six possible values for race: C=White, M=Asian or Pacific Islander, N=Black, R=American Indian or Alaskan Native, X=Other and Z=Unknown.

There are no missing values for race. However, approximately one half of the records are classified as Unknown. These records are distributed across all MTFs, indicating that the

problem is not site- or CHCS server-specific. This may limit the ability to use the data to look at diseases or conditions that disproportionately affect one race unless data are used in combination with a supplemental database to obtain demographic information.

Record Type

DHSS enters a value of "POP" for all Pharmacy Outpatient records. The value does not change between extracts and there are no missing records. The field has an analytic value if different HL7 datasets are used in a single analysis as it will serve to identify the source of the individual record.

Requesting DMIS Facility Name

This field is the text translation of the DMIS ID provided in the Requesting DMIS ID field. This field is assigned by DHSS at the request of the EDC. This assignment is requested because the translation of the DMIS code on the official list is often more accurate than the Requesting Facility field in CHCS. This allows for more accurate investigations when geographic information is used. Because this field is a translation of the Requesting Facility field in CHCS, it will be missing when that field is missing in the record.

Requesting DMIS ID

The Requesting DMIS ID is a four digit code assigned by the DOD to all units within all of the installations to uniquely identify them. This field is included by DHSS at the request of the EDC. The EDC provided an official DMIS list to DHSS for this purpose. DHSS translated the Requesting Facility field within CHCS to an assigned DMIS code. This code allows MTFs to be grouped based on geographic location, and parent/child relationships between installations to be identified. Since this field is calculated based on the Requesting Facility field, all records missing a value for that field will be missing a value for the Requesting DMIS ID field. Since several codes begin with zeros they are often dropped when importing the data from a text file, which may lead to problems when trying to produce summary statistics based on MTF.

Requesting Facility

The Requesting Facility field is the field in CHCS that indicates the name of the MTF where the order originated, and is a relatively standard text field. Problems may be encountered if the text is entered incorrectly when the facility is registered in the system (e.g. misspellings). The field allows tracking of orders from their origin to where they are filled.

Requesting Facility Service

The Requesting Facility Service field indicates the branch of service that the MTF is associated with. This value is determined from the DMIS code list provided to DHSS by the EDC. It will be missing from a record when the Requesting Facility information is missing. The possible values are: A=Army, C=Coast Guard, F=Air Force, N=Navy. This field is useful for limiting the observations by service. Often, the data to which the HL7 Pharmacy data will be compared to are limited by branch of service for the MTF or the patient. If this is the case, the HL7 Pharmacy data can be limited to the same parameters.

Requesting Work Center

The Requesting Work Center is the ward or clinic within the MTF that requests the prescription. This field is an unstructured text field with many possible values. Many medications have multiple indications that can lead to epidemiologically different assumptions. This field may help clarify the reason behind the prescription.

Service

The service field refers to the branch of the sponsor. The value is determined from the first component of the PatCat field and the values are the same. Therefore, there are 902 records that are missing the Branch of Service, which is the same as the number of missing PatCat codes. The distribution of values for service reflects the overall population distribution among the branches.

Sponsor ID

The Sponsor ID field corresponds to the Social Security Number (SSN) of the sponsor and is in the format of xxxxxxxx, with no dashes. This field is relatively complete with only two records from extracts II and III missing values. Those two records are the same as those that are missing an FMP value.

Sponsor ID is not sufficient to serve as a unique identifier for each patient but it can be used in conjunction with the FMP to identify each patient. Creating a unique identifier is necessary because the patient's SSN is not included consistently in any of the HL7 Pharmacy data strings. It is important to preserve the entire Sponsor ID or SSN when importing the data into any analysis program. If the field is not properly coded as a character field, leading zeros can be dropped.

Sponsor UIC Code

Unit Identification Code (UIC) is a five digit number that identifies the assigned duty station of the Sponsor. The UIC serves as a way to group patients geographically according to where the Sponsor is assigned, and allows for the identification of geographic clusters. Note that not all dependents reside in the same geographic location as the Sponsor, especially when the service members are assigned to ships.
This field remains unpopulated but is included in the final daily feeds.

Sponsor UIC Description

This field is a text translation of the Sponsor UIC Code field. The translation allows researchers to identify the location description of the UIC code without having to use a directory. This field remains unpopulated but is included in the final daily feeds.

Time of Most Recent Refill

This field represents the time component of the Date of Most Recent Refill field, formatted: HHMM.

Time of Transaction

The field represents the time component of the Date of Transaction, formatted; HHMM.
Units

Units

The Units field indicates the unit of medication that is dispensed and relates to the Amount field. This is a standard text field with possible values of ML, TUBE, INH (inhaler), VIAL, BTL (bottle), PK (pack), JAR, TAB, SYR (syringe) and several others. Not all values are translated in the data dictionary.

Almost half of records are missing a Units value. The units can be determined by searching an NDC code directory for most medications. There are still some formulations that do not contain this information.

Appendix A: Examples of projects using Pharmacy OP data

HL7 Pharmacy OP and Lab Correlation Study
Ashleigh K. McCabe, MPH, A. Kate Goodin, MPH, Gosia Kubiak, MSc, MPH, Asha Riegodedios, MSPH, Karen Otero-Fisher, Craig Nash

Background: The Navy and Marine Corps Public Health Center receives Health Level 7 (HL7) Pharmacy and Laboratory data daily. A correlation study between the Pharmacy and Laboratory data was pursued to understand the relationship between the datasets as it pertains to influenza case identification.

Methods: Influenza medication prescriptions for the 2006-2007 influenza season were extracted from the HL7 Pharmacy data and were compared to the Laboratory records during the same period. Pharmacy data were also matched to ambulatory care data to determine if treatment was based on physician diagnosis.

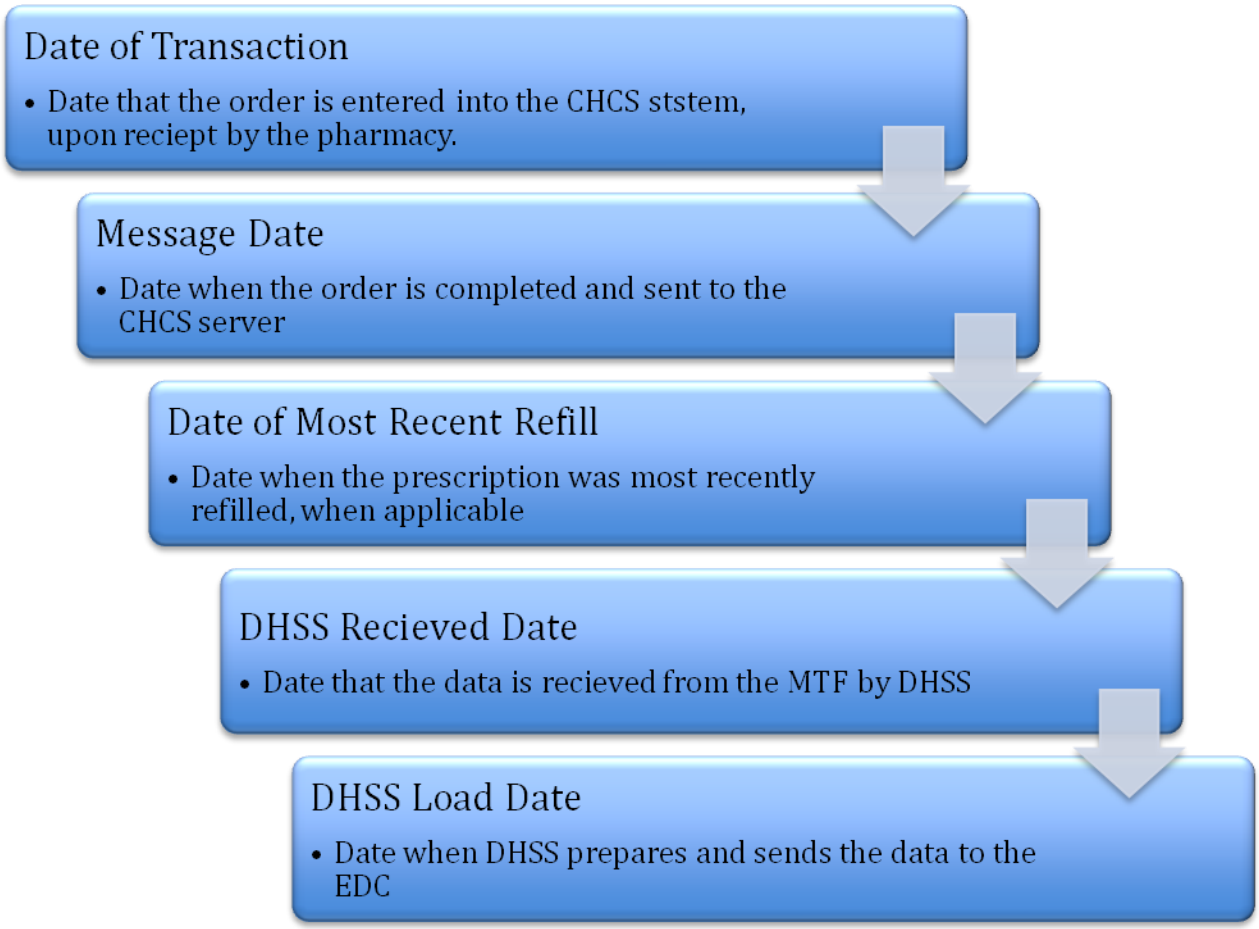
Influenza Antivirals: The table below describes the four influenza-specific antiviral medications. Identification of these medications in the data was used to identify cases that may or may not have been laboratory tested and will allow for broader case capture. Dosage for each drug depends on age, weight and concurrent medical conditions. Although amantadine can also be used to treat Parkinson's disease and other movement disorders, doses for treatment of influenza are much lower; as such, dose information can be used to differentiate possible diagnoses.

Medication	Route	Strand
Amantadine	Oral (tablet or syrup)	A
Oseltamivir	Oral (capsule or suspension)	A & B
Rimantadine	Oral (tablet or syrup)	A
Zanamivir	Inhaled	A & B

Results: A substantial number of influenza-specific pharmacy records did not have corresponding laboratory or ambulatory care records. Sixty percent of all prescriptions did not have an associated influenza test.

Conclusions: Many influenza patients were captured in the Pharmacy data that would not have met the Laboratory data case definition. Though 90% of patients that were tested had positive results, there are also a significant proportion of cases that were treated with influenza medications based on solely symptomology.

Appendix B: Timeline of Dates



* The timeline above represents the relationship of the dates that is most often observed in the HL7 Pharmacy Outpatient dataset. Some records do not follow this timeline due to the timeliness of data entry.

Abbreviations and Acronyms

CHCS	Composite Health Care System
DEERS	Defense Eligibility Enrollment Reporting System
DHSS	Defense Health Services System
DMIS	Defense Medical Information System
DOD	Department of Defense
EDC	EpiData Center
FDA	Federal Drug Administration
FMP	Family Member Prefix
HL7	Health Level 7
ICD-9-CM	International Classification of Diseases, 9th Revision, Clinical Modification
MEPRS	Medical Expense and Performance Reporting System
MHS	Military Health System
MTF	Military Treatment Facility
NDC	National Drug Code (Can be in the form of a code or a name)
NMCPHC	Navy and Marine Corps Public Health Center
OP	Outpatient (Pharmacy Dataset)
PATCAT	Patient Category Code
PDTS	Pharmacy Data Transaction Service
SADR	Standard Ambulatory Data Record
SIDR	Standard Inpatient Data Record
SSN	Social Security Number
UIC	Unit Identification Code

Appendix C: Glossary

True Duplicates	Records in which all fields are exactly the same; only one of these records should be used in analysis.
Unique Order	All records associated with each specific drug prescription.
Unique Record	All transactions associated with each prescription for an individual patient.
Unique Patient	The combination of Sponsor ID and FMP creates a unique identifier that can be used to track individual patients through all pharmacy records.

POINT OF CONTACT

Navy and Marine Corps Public Health Center
Communicable Disease Division
EpiData Center Department
WWW.NMCPHC.MED.NAVY.MIL/

Ashleigh K. McCabe
757.953.0692
ashleigh.mccabe@med.navy.mil